hep-ph/0002032
FERMILAB-Pub-00/032-T
CTEQ-015
MSUHEP-00126
UH-511-954-00

Higgs Production:

A Comparison of Parton Showers and Resummation

C. Balázs[1*], J. Huston[2†], and I. Puljak[3‡]

[1] *Department of Physics & Astronomy, University of Hawaii, Honolulu, HI 96822, USA*

[2] *Department of Physics & Astronomy, Michigan State University, East Lansing, MI 48824, USA*

[3] *Laboratoire de Physique Nucléaire et des Hautes Energies, Ecole Polytechnique, 91128 Palaiseau, France, and University of Split, 21000 Split, Croatia*

(February 2, 2000)

The search for the Higgs boson(s) is one of the major priorities at the upgraded Fermilab Tevatron and at the CERN Large Hadron Collider (LHC). Since Monte Carlo event generators are heavily utilized to understand and extract any possible Higgs signal, it is crucial to establish the reliability of their predictions. An understanding of the Higgs signature depends upon the understanding of the details of the soft-gluon emission from the initial state partons. In this paper, the parton shower formalism for Monte Carlos is compared to that of an analytic resummation calculation. Predictions and, where they exist, data for the transverse momentum distribution of Higgs bosons, Z^0 bosons, and photon pairs are compared for the Tevatron and the LHC. This comparison is useful in understanding the strengths and the weaknesses of the different theoretical approaches, and in testing their reliability.

PACS numbers: 12.38.Cy, 14.80.Bn, 13.85.Qk, 12.20.Fv.

*balazs@phys.hawaii.edu

†huston@pa.msu.edu

‡Ivica.Puljak@cern.ch

I. INTRODUCTION

To reveal the dynamics of electroweak symmetry breaking, a new generation of hadron colliders will search for the Higgs boson(s). The potential of the upgraded Fermilab Tevatron, which starts operation within a year, was analysed in Ref. [1]. Later in this decade, two experimental collaborations (ATLAS and CMS) join the search at the CERN Large Hadron Collider (LHC) with 14 TeV center of mass proton-proton collisions. An extraction of the Higgs signal at the LHC requires not only the precise knowledge of the signal and background invariant mass distributions, but also the accurate prediction of the corresponding transverse momentum (p_T) distributions. In general, the determination of the signal requires a detailed event modeling, an understanding of the detector resolution, kinematical acceptance, and efficiency, all of which depend on the p_T distribution. The p_T distribution can also be used to devise an improved strategy for the Higgs search, and to enhance the statistical significance of the signal over the background [2,3]. In the $gg \to HX \to \gamma\gamma X$ mode at the LHC, for example, the shape of the signal and the background p_T distribution of the photon pairs is different (c.f. Refs. [4,5]), with the signal being harder. This difference can be utilized to increase the signal to background ratio. Furthermore, since vertex pointing with the photons is not possible in the CMS barrel, the shape of the p_T distribution affects the precision of the determination of the event vertex from which the Higgs (decaying into two photons) originated. [1] Thus, for a successful, high precision extraction of the Higgs signal, the theoretical calculation must be capable of reproducing the expected transverse momentum distribution.

To reliably predict the p_T distribution of Higgs bosons at the LHC, especially for the low to medium p_T region where the bulk of the rate is, the effects of the multiple soft-

[1]The vertex with the most activity is chosen as the vertex from which the Higgs particle has originated. If the Higgs is typically produced at a relatively high value of p_T, then this choice is correct a large fraction of the time.

gluon emission have to be included. One approach which achieves this is parton showering [6]. Parton shower Monte Carlo programs such as PYTHIA [7], HERWIG [8] and ISAJET [9] are commonly used by experimentalists, both as a way of comparing experimental data to theoretical predictions, and also as a means of simulating experimental signatures in kinematic regimes for which there are no experimental data yet (such as for the LHC). The final output of these Monte Carlo programs consists of the 4-momenta of a set of final state particles. This output can either be compared to reconstructed experimental quantities or, when coupled with a simulation of a detector response, can be directly compared to raw data taken by the experiment, and/or passed through the same reconstruction procedures as the raw data. In this way, the parton shower programs can be more useful to experimentalists than analytic calculations. Indeed, almost all of the physics plots in the ATLAS physics TDR [10] involve comparisons to PYTHIA (version 5.7).

Predictions of the Higgs p_T can also be obtained utilizing analytic resummation formalisms, which sum contributions of $\alpha_s^m(Q)\ln^n(Q/p_T)$ (where Q is the Higgs invariant mass) up to all orders in the strong coupling α_s. In the recent literature, most calculations of this kind are either based on, or originate from, the formalism developed by J.C. Collins, D.E. Soper, and G. Sterman (CSS) [11] (for the latest review see Ref. [12]). The CSS formalism takes into account the effects of the multiple soft-gluon emission while reproducing the rate, systematically including the higher order QCD corrections. It is possible to smoothly match the CSS result to the fixed order one in the medium to high p_T region, thus obtaining a prediction for the full p_T distribution [13]. In this paper, we use this formalism as the analytic 'benchmark' to calculate the p_T distributions of Higgs bosons at the LHC, and of Z^0 bosons and photon pairs produced in hadron collisions.

For many physical quantities, the predictions from parton shower Monte Carlo programs should be nearly as precise as those from analytic theoretical calculations. It is expected that both the Monte Carlo and analytic calculations should accurately describe the effects of the emission of multiple soft-gluons from the incoming partons, an all orders problem in QCD. The initial state soft-gluon emission affects the kinematics of the final state partons. This

may have an impact on the signatures of physics processes at both the trigger and analysis levels and thus it is important to understand the reliability of such predictions. The best method for testing the reliability is a direct comparison of the predictions to experimental data. If no experimental data are available for certain predictions, then some understanding of the reliability may be gained from the comparison of the predictions from the two different methods.

Production of a light, neutral Higgs boson at the LHC in the standard model and its supersymmetric extensions proceeds via the partonic subprocess gg (through heavy fermion loop) $\to HX$ [14]. One of the major backgrounds for a light Higgs, in the mass range of 100 GeV $\lesssim m_H \lesssim$ 150 GeV, is diphoton production, a sizable contribution to which comes from the same, gg initial state. Since the major part of the soft-gluon radiation is initiated from the incoming partons, the structures of the resummed corrections are similar for Higgs boson and diphoton production. Because the latter is measurable at the Fermilab Tevatron, diphoton production provides an exceptional opportunity to test the different theoretical models. Z^0 boson production can also be a good testing ground for the soft-gluon corrections to Higgs production. The treatment of the fixed order and resummed QCD corrections is theoretically well understood and implemented at next-to-next-to-leading order [13]. Furthermore, just like in the diphoton case, predictions can also be tested against Tevatron data.

II. PARTON SHOWERING AND RESUMMATION

For technical reasons, the initial state parton shower proceeds by a *backwards* evolution, starting at the large (negative) Q^2 scale of the hard scatter and then considering emissions at lower and lower (negative) virtualities, corresponding to earlier points on the cascade (and earlier points in time), until a scale corresponding to the factorization scale is reached. The transverse momentum of the initial state is built up from the whole series of splittings (and boosts). The showering process is independent of the hard scattering process being

considered (as long as one does not introduce any matrix element corrections), and depends only on the initial state partons and the hard scale of the process.

Parton showering utilizes the fact that the leading order collinear singularities of cross sections factorize in the collinear limit. This is expressed mathematically as

$$\lim_{p_g \to p_b} |\mathcal{M}_{n+1}|^2 = g_s^2 (p_b . p_g)^{-1} P_{g \leftarrow a}(z) |\mathcal{M}_n|^2, \tag{1}$$

where $\mathcal{M}_{n+1}$ is the invariant amplitude for the process producing n partons and a gluon, g_s is the strong coupling constant, p_b and p_g are the 4-momenta of the daughters of parton a, and $P_{g \leftarrow a}(z)$ is the DGLAP evolution kernel associated with the $a \to bg$ splitting. These leading order collinear singularities can be factorized into a Sudakov form factor: $\mathcal{S}_{shower} = 1 - \mathcal{P}(\text{no emission}) = exp\{-\int dp^2/p^2 \int dz P(z)\}$, where $\mathcal{P}$ denotes probability. The distribution $1 - \mathcal{S}_{shower}$ can be used to generate the Q^2 for the first emission and hence for the whole cascade. The formalism can be extended to soft singularities as well by using angular ordering. In this approach, the choice of the hard scattering is based on the use of evolved parton distributions, which means that the inclusive effects of initial-state radiation are already included. What remains is, therefore, to construct the exclusive showers.

Parton showering resums primarily the leading logs, which are universal, i.e. process independent, and depend only on the given initial state. In this lies one of the strengths of Monte Carlos, since parton showering can be incorporated into a wide variety of physical processes. An analytic calculation, in comparison, can resum all logs. For example, the CSS formalism sums all of the logarithms with Q^2/p_T^2 in their arguments, where (for Higgs boson production) Q is the 4-momentum of the Higgs and p_T is its transverse momentum. As discussed in Refs. [12,13], all of the 'dangerous logs' are included in the Sudakov exponent, which can be written in the impact parameter (b) space as:

$$\mathcal{S}(Q, b) = \int_{C_0^2/b^2}^{Q^2} \frac{d\overline{\mu}^2}{\overline{\mu}^2} \left[A\left(\alpha_s(\overline{\mu})\right) \ln\left(\frac{Q^2}{\overline{\mu}^2}\right) + B\left(\alpha_s(\overline{\mu})\right) \right],$$

with the A and B functions being free of large logs and perturbatively calculable:

$$A\left(\alpha_s(\bar{\mu})\right) = \sum_{n=1}^{\infty} \left(\frac{\alpha_s(\bar{\mu})}{\pi}\right)^n A^{(n)}, \quad B\left(\alpha_s(\bar{\mu})\right) = \sum_{n=1}^{\infty} \left(\frac{\alpha_s(\bar{\mu})}{\pi}\right)^n B^{(n)},$$

and $C_0 = 2e^{-\gamma_{Euler}}$.

These functions contain an infinite number of coefficients, with the $A^{(n)}$ coefficients being universal, while the $B^{(n)}$'s are process dependent. In practice, the number of towers of logarithms included in the Sudakov exponent depends on the level to which a fixed order calculation was performed for a given process. For example, if only a next-to-leading order (NLO) calculation is available, only the coefficients $A^{(1)}$ and $B^{(1)}$ can be included. If a next-to-next-to-leading order calculation is available, then $A^{(2)}$ and $B^{(2)}$ can be extracted and incorporated into a resummation calculation, and so on. This is the case, for example, for Z^0 boson production. So far, only the $A^{(1)}$, $A^{(2)}$ and $B^{(1)}$ coefficients are known for Higgs production but the calculation of $B^{(2)}$ is in progress [15]. If we try to interpret parton showering in the same language, which is admittedly risky, then we can say that the Monte Carlo Sudakov exponent always contains a term analogous to $A^{(1)}$. It was shown in Reference [16] that a suitable modification of the Altarelli-Parisi splitting function, or equivalently the strong coupling constant α_s, also effectively approximates the $A^{(2)}$ coefficient. [2]

In contrast with the shower Monte Carlos, analytic resummation calculations integrate over the kinematics of the soft-gluon emission, with the result that they are limited in their predictive power for inclusive final states. While the Monte Carlo maintains an exact treatment of the branching kinematics, in the original CSS formalism no kinematic penalty is paid for the emission of the soft-gluons, although an approximate treatment of this can be incorporated into its numerical implementations, such as ResBos [5,13]. Neither the parton showering process nor the analytic resummation translate smoothly into kinematic configurations where one hard parton is emitted (at large p_T). In the Monte Carlo matrix element corrections, and in the analytic resummation calculation, matching is necessary. This matching is standard procedure for resummation calculations, and matrix element corrections are becoming increasingly common in Monte Carlos [17–19].

---

[2]This is rigorously true only for the high x or $\sqrt{\tau}$ region.

With the appropriate input from higher order cross sections, a resummation calculation has the corresponding higher order normalization and scale dependence. The normalization and scale dependence for the Monte Carlo, though, remains that of a leading order calculation. The parton showering process redistributes the event particles in phase space, but does not change the total cross section (for example, for the production of a Higgs boson). [3]

One quantity which should be well described by both calculations is the transverse momentum (p_T) of the final state electroweak boson in a subprocess such as $q\overline{q} \to W^{\pm}X$, Z^0X or $gg \to HX$, where most of the p_T is provided by initial state parton emission. The parton showering supplies the same sort of transverse kick as the soft-gluon radiation in a resummation calculation. Indeed, very similar Sudakov form factors appear in both approaches, with the caveats about the $A^{(n)}$ and $B^{(n)}$ terms mentioned previously. This correspondence between the Sudakov form factors in resummation and Monte Carlo approaches may seem trivial, but there are many subtleties between the two approaches relating to both the arguments of the Sudakov factors as well as the impact of sub-leading logs [20].

At a point in its evolution corresponding to (typically) the virtuality of a few GeV^2, the parton shower is cut off and the effects of gluon emission at softer scales must be parameterized and inserted by hand. This is similar to the (somewhat arbitrary) division between perturbative and non-perturbative regions in a resummation calculation. The parameterization is typically expressed in a Gaussian form, similar to that used for the non-perturbative k_T in a resummation program. In general, the value for the non-perturbative $\langle k_T \rangle$ needed in a Monte Carlo program will depend on the particular kinematics being investigated. In the case of the resummation calculation, the non-perturbative physics is determined from fits to fixed target data and then automatically evolved to the kinematic regime of interest.

---

[3] Technically, one could add the branching for $q \to q$+Higgs in the shower, which would somewhat increase the Higgs cross section. However, the main contribution to the higher order K-factor comes from the virtual corrections and the 'Higgs bremsstrahlung' contribution is negligible.

A value for the average non-perturbative k_T of greater than 1 GeV does not imply that there is an anomalous intrinsic k_T associated with the parton size. Rather, this amount of $\langle k_T \rangle$ needs to be supplied to provide what is missing in the truncated parton shower. If the shower is cut off at a higher virtuality, more of the 'non-perturbative' k_T will be needed.

III. Z^0 BOSON PRODUCTION AT THE TEVATRON

The 4-vector of a Z^0 boson, and thus its transverse momentum, can be measured with great precision in the e^+e^- decay mode. Resolution effects are relatively minor and are easily corrected. Thus, the Z^0 p_T distribution is a great testing ground for both the resummation and Monte Carlo formalisms for soft-gluon emission. The (resolution corrected) p_T distribution for Z^0 bosons (in the low p_T region) for the CDF experiment [21] is shown in Figure 1, compared to both the resummed prediction from ResBos, and to two predictions from `PYTHIA` (version 6.125). One `PYTHIA` prediction uses the default (rms)[4] value of intrinsic k_T of 0.44 GeV and the second a value of 2.15 GeV (per incoming parton). [5] The latter value was found to give the best agreement for `PYTHIA` with the data. [6] All of the predictions use the CTEQ4M parton distributions [23]. The shift between the two `PYTHIA` predictions at low p_T is clearly evident. As might have been expected, the high p_T region (above 10 GeV) is unaffected by the value of the non-perturbative k_T. Note that much of the k_T 'given' to the incoming partons at their lowest virtuality, Q_0, is reduced at the hard scatter due to the number of gluon branchings preceding the collision. The emitted gluons carry off a sizable

---

[4] For a Gaussian distribution, $k_T^{rms} = 1.13\langle k_T \rangle$.

[5] A previous publication [17] indicated the need for a substantially larger non-perturbative $\langle k_T \rangle$, of the order of 4 GeV, for the case of $W^{\pm}$ production at the Tevatron. The data used in the comparison, however, were not corrected for resolution smearing, a fairly large effect for the case of $W \to e\nu$ production and decay.

[6] A similar conclusion has been reached for comparisons of the CDF Z^0 p_T data with `HERWIG` [22].

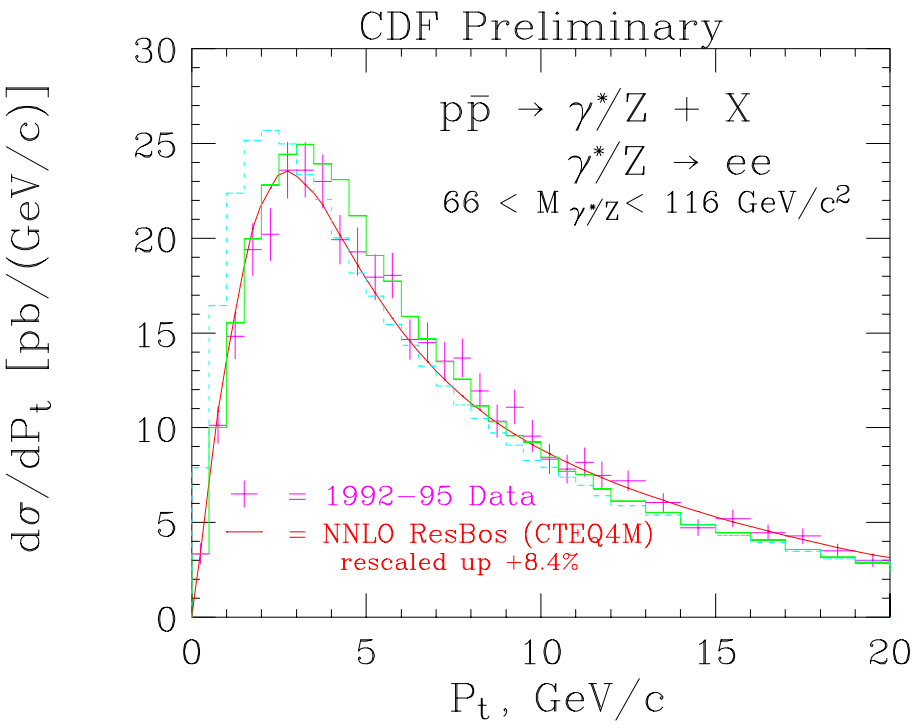

FIG. 1. The Z^0 p_T distribution (at low p_T) from CDF for Run 1 compared to predictions from ResBos (curve) and from PYTHIA (histograms). The two PYTHIA predictions use the default (rms) value for the non-perturbative k_T (0.44 GeV) and the value that gives the best agreement with the shape of the data (2.15 GeV).

fraction of the original non-perturbative k_T. This point will be investigated in more detail later for the case of Higgs production.

As an exercise, one can transform the resummation formula in order to bring it to a form where the non-perturbative function acts as a Gaussian type smearing term. Using the Ladinsky-Yuan parameterization [24] of the non-perturbative function in ResBos leads to an rms value for the effective k_T smearing parameter, for Z^0 production at the Tevatron, of 2.5 GeV. This is similar to that needed for PYTHIA and HERWIG to describe the Z^0 production data at the Tevatron.

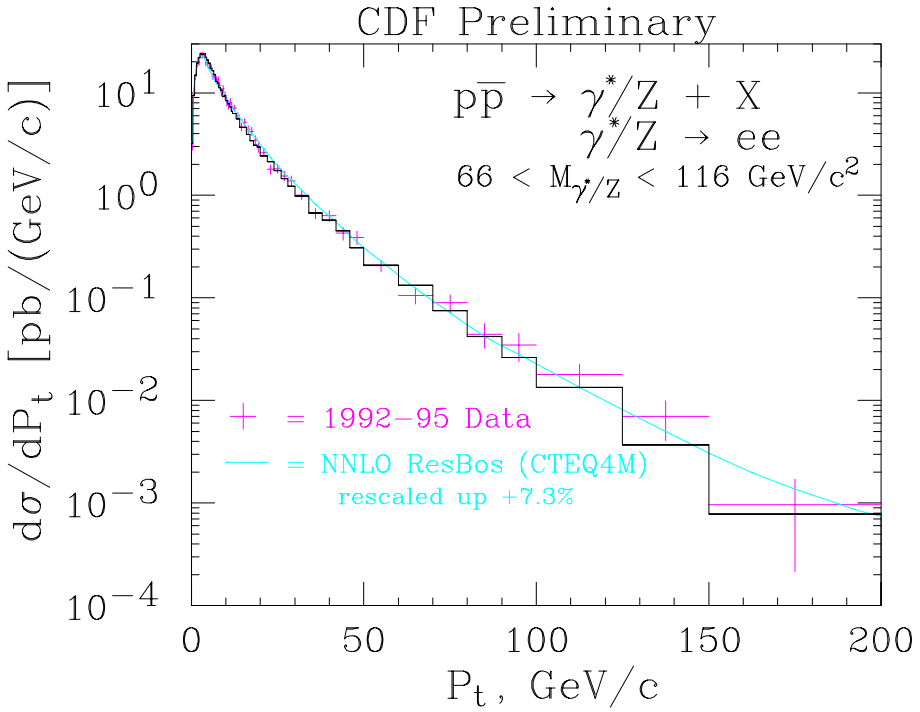

FIG. 2. The Z^0 p_T distribution (for the full range of p_T) from CDF for Run 1 compared to predictions from ResBos (curve) and from PYTHIA (histogram).

In Figure 1, the normalization of the resummed prediction was rescaled upwards by 8.4%. The PYTHIA prediction was rescaled by a factor of 1.3-1.4 (remember that this is only a leading order comparison) for the shape comparison.

As stated previously, the resummed prediction correctly describes the shape of the Z^0 p_T distribution at low p_T but, even with the optimal non-perturbative k_T, there is still a noticeable shape difference between the Monte Carlo and resummed prediction. It is interesting to note that if the process dependent coefficients ($B^{(1)}$ and $B^{(2)}$) were not incorporated into the resummation prediction, the result would be an increase in the height of the peak and a decrease in the rate between 10 and 20 GeV, leading to a better agreement with the PYTHIA prediction [12].

The Z^0 p_T distribution is shown over a wide p_T range in Figure 2. The PYTHIA and ResBos predictions both describe the data well. Note especially the agreement of PYTHIA with the data at high p_T, made possible by explicit matrix element corrections (from the subprocesses $q\overline{q} \to Z^0 g$ and $gq \to Z^0 q$) to the Z^0 production process. [7]

IV. DIPHOTON PRODUCTION

Most of the experience that we have for comparisons of data to resummation calculations/Monte Carlos deals with Drell-Yan production, i.e. $q\overline{q}$ initial states. It is important then to examine diphoton production at the Tevatron, where a large fraction of the contribution at low mass is due to gg scattering. The prediction for the diphoton p_T distribution at the Tevatron, from PYTHIA (version 6.122), is shown in Figure 3, using the experimental cuts applied in the CDF analysis [25]. It is interesting to note that about half of the diphoton cross section at the Tevatron is due to the gg subprocess, and that the diphoton p_T distribution is noticeably broader for the gg subprocess than for the $q\overline{q}$ subprocess.

A comparison of the p_T distributions for the two diphoton subprocesses ($q\overline{q}, gg$) in two recent versions of PYTHIA, 5.7 and 6.1, is shown in Figure 4. There seems to be little difference in the p_T distributions between the two versions for both subprocesses. As will be shown later, this is not true for the case of Higgs production.

---

[7] Slightly different techniques are used for the matrix element corrections by PYTHIA [17] and by HERWIG [18]. In PYTHIA, the parton shower probability distribution is applied over the whole phase space and the exact matrix element corrections are applied only to the branching closest to the hard scatter. In HERWIG, the corrections are generated separately for the regions of phase space unpopulated by HERWIG (the 'dead zone') and the populated region. In the dead zone, the radiation is generated according to a distribution using the first order matrix element calculation, while the algorithm for the already populated region applies matrix element corrections whenever a branching is capable of being 'the hardest so far'.

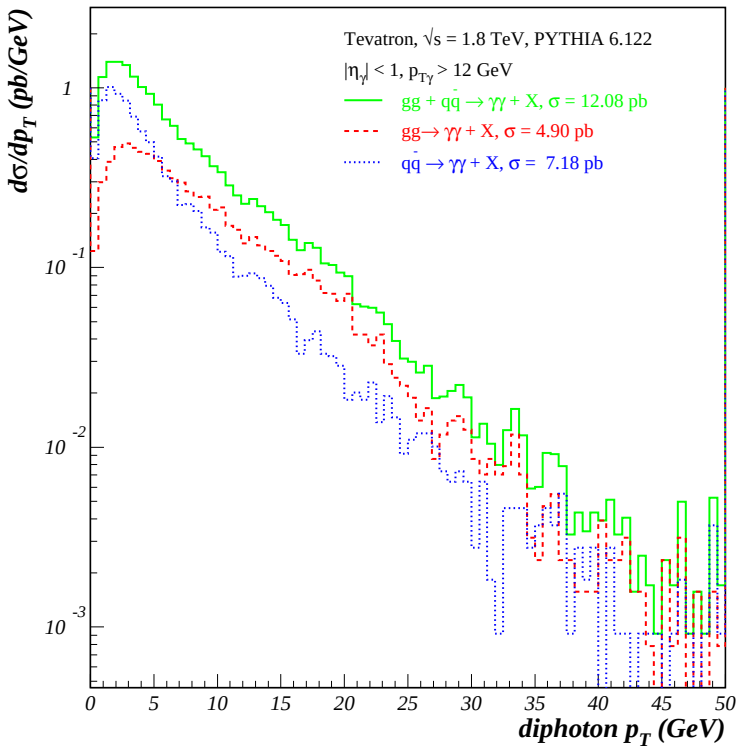

FIG. 3. A comparison of the PYTHIA predictions for diphoton production at the Tevatron for the two different subprocesses, $q\overline{q}$ and gg. The same cuts are applied to PYTHIA as in the CDF diphoton analysis.

In Figure 5 are shown the ResBos predictions for diphoton production at the Tevatron from $q\overline{q}$ and gg scattering compared to the PYTHIA predictions (using the same experimental cuts). The gg subprocess predictions in ResBos agree well with those from PYTHIA while the $q\overline{q}$ p_T distribution is noticeably broader in ResBos. The latter behavior is due to the presence of the Y piece in ResBos at moderate p_T, and the matching of the $q\overline{q}$ cross section to the fixed order $q\overline{q} \to \gamma\gamma g$ at high p_T. The corresponding matrix element correction is not in PYTHIA. It is interesting to note that the PYTHIA and ResBos predictions for $gg \to \gamma\gamma$ agree in the moderate p_T region, even though the ResBos prediction has the Y piece present and is matched to the matrix element piece $gg \to \gamma\gamma g$ at high p_T, while there is no such matrix element correction for PYTHIA. This shows the smallness of the Y piece for the gg subprocess, which is the same conclusion that was reached in Ref. [3]. One way to understand this is

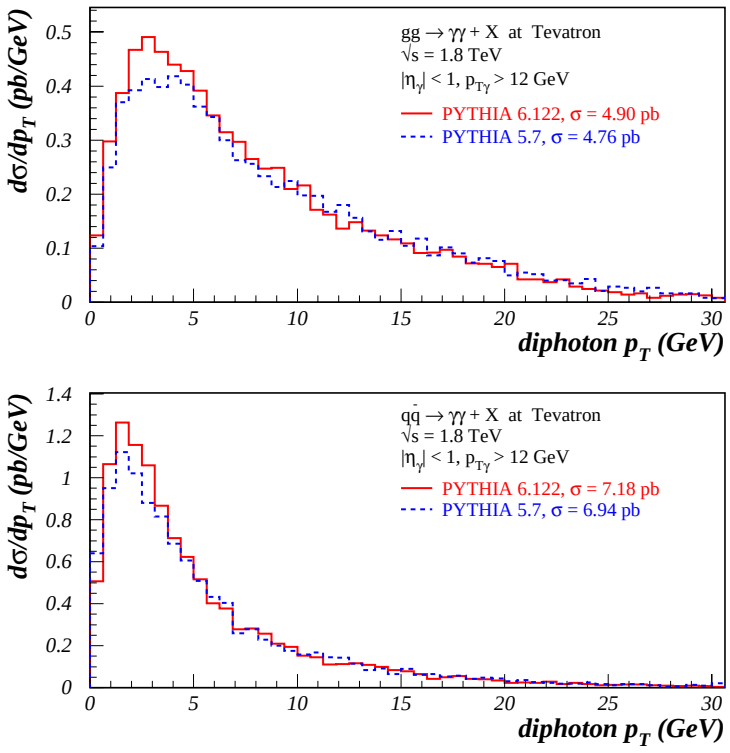

FIG. 4. A comparison of the PYTHIA predictions for diphoton production at the Tevatron for the two different subprocesses, gg (top) and $q\overline{q}$ (bottom), for two recent versions of PYTHIA. The same cuts are applied to PYTHIA as in the CDF diphoton analysis.

recalling that the gg parton-parton luminosity falls very steeply with increasing partonic center of mass energy, $\sqrt{\hat{s}}$. This falloff tends to suppress the size of the Y piece since the production of the diphoton pair at higher p_T requires larger x_1, x_2 values. In the default CSS formalism, there is no such kinematic penalty in the resummed piece since the soft-gluon radiation comes for "free". (Larger x_1 and x_2 values are not required.)

Comparisons of the diphoton data measured by both the CDF [25] and D0 [26] experiments indicate a disagreement of the observed diphoton p_T distribution with the NLO QCD predictions [27]. In particular, the p_T distribution in the data is noticeably broader than that predicted by fixed order QCD calculations, but in agreement with the predictions of ResBos. The transverse momentum distribution of the diphoton pair is particularly sensitive to the effects of soft-gluon radiation, since the distribution is a delta function at leading

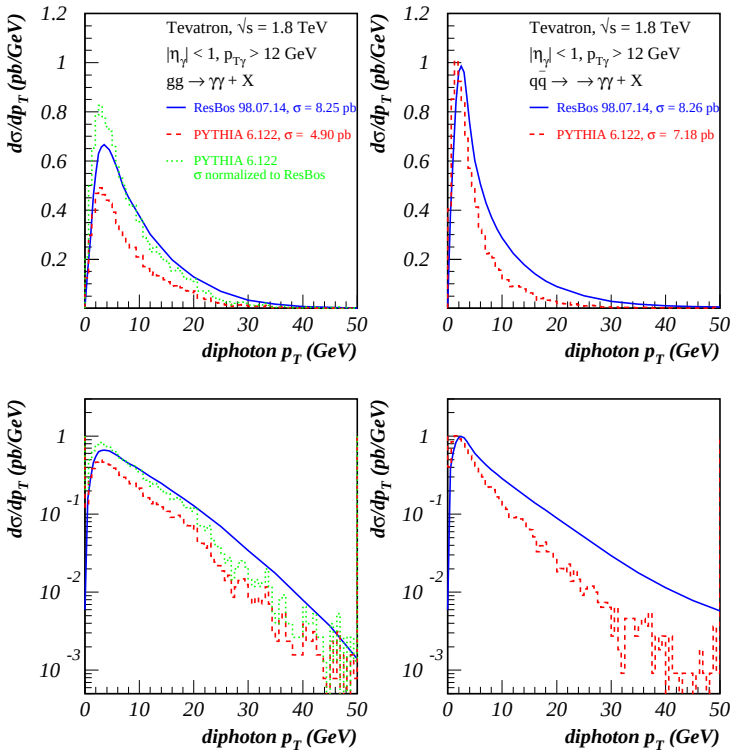

FIG. 5. A comparison of the PYTHIA and ResBos predictions for diphoton production at the Tevatron for the two different subprocesses, gg (left) and $q\overline{q}$ (right). The same cuts are applied to PYTHIA and ResBos as in the CDF diphoton analysis. The bottom figures show the same in logarithmic scale.

order. Given the small size of the diphoton cross section at the Tevatron, the comparisons for Run 1 are statistically limited. A more precise comparison with the effects of soft-gluon radiation will be possible with the 2 fb^{-1} or greater data sample that is expected for CDF and D0 in Run 2.

The prediction for the diphoton production cross section, as a function of the diphoton p_T and using cuts appropriate to ATLAS and CMS, is shown in Figure 6. Note that, as at the Tevatron, about half of the cross section is due to gg scattering and the diphoton p_T distribution from gg scattering is noticeably broader than that from $q\overline{q}$ production.

In Figure 7 is shown a comparison of the diphoton p_T distribution for the two different versions of PYTHIA, for the two different subprocesses. Note that the p_T distribution in PYTHIA version 5.7 is somewhat broader than that in version 6.122 for the case of gg

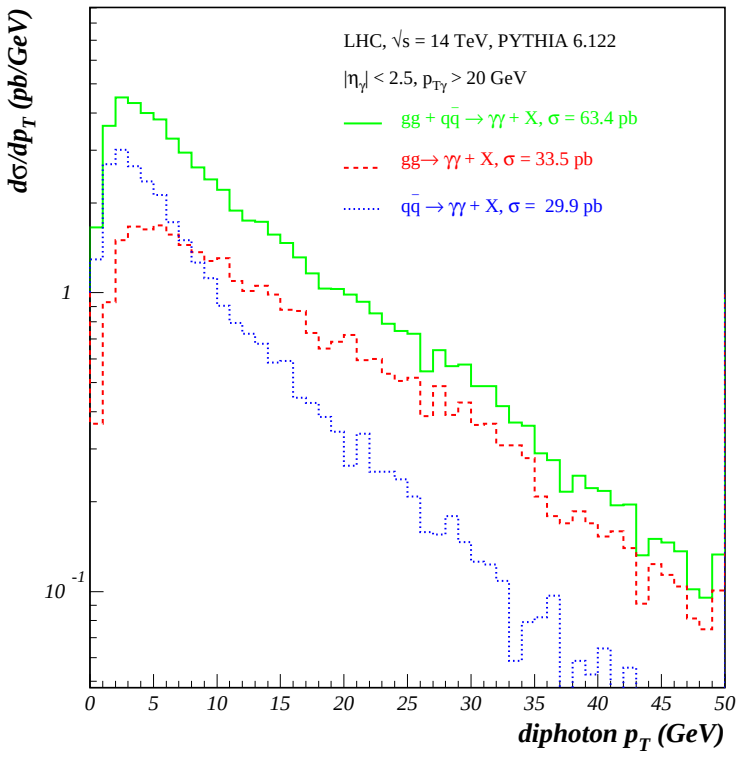

FIG. 6. A comparison of the PYTHIA predictions for diphoton production at the LHC for the two different subprocesses, $q\overline{q}$ and gg. Similar cuts are applied to the diphoton kinematics as those used by ATLAS and CMS.

scattering. The effective diphoton mass range being considered here is lower than the 150 GeV Higgs mass that will be considered in the next section. As will be seen, the differences in soft-gluon emission between the two versions of PYTHIA are larger in that case.

In Figure 8 are shown the ResBos predictions for diphoton production at the LHC from $q\overline{q}$ and gg scattering compared to the PYTHIA predictions (using the same experimental cuts). Again, the gg subprocess predictions in ResBos agree well with those from PYTHIA, while the $q\overline{q}$ p_T distribution is noticeably broader in ResBos, for the reasons cited previously.

V. HIGGS BOSON PRODUCTION

A comparison of the Higgs p_T distribution at the LHC, for a Higgs mass of 150 GeV, is shown in Figure 9, for ResBos and the two recent versions of PYTHIA. As before, PYTHIA

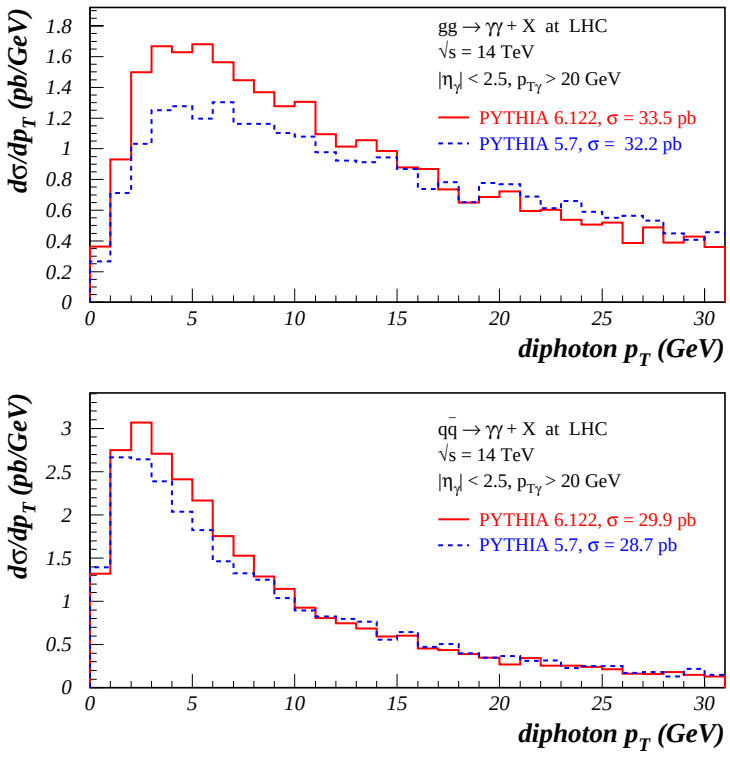

FIG. 7. A comparison of the PYTHIA predictions for diphoton production at the LHC for the two different subprocesses, gg (top) and $q\overline{q}$ (bottom), for two recent versions of PYTHIA. Similar cuts are applied to the diphoton kinematics as are used by ATLAS and CMS.

has been rescaled to agree with the normalization of ResBos to allow for a better shape comparison. Note that the peak of the resummed distribution has moved to $p_T \approx 11$ GeV (compared to about 3 GeV for Z^0 production at the Tevatron). This is partially due to the larger mass (150 GeV compared to 90 GeV), but is primarily because of the larger color factors associated with initial state gluons ($C_A = 3$) rather than quarks ($C_F = 4/3$), and also because of the larger phase space for initial state gluon emission at the LHC.

The newer version of PYTHIA agrees well with ResBos at low to moderate p_T, but falls below the resummed prediction at high p_T. This is easily understood: ResBos switches to the NLO Higgs + jet matrix element at high p_T while the default PYTHIA can generate the Higgs p_T distribution only by initial state gluon radiation, using as maximum virtuality the Higgs mass squared. High p_T Higgs production is another example where a $2 \to 1$ Monte Carlo

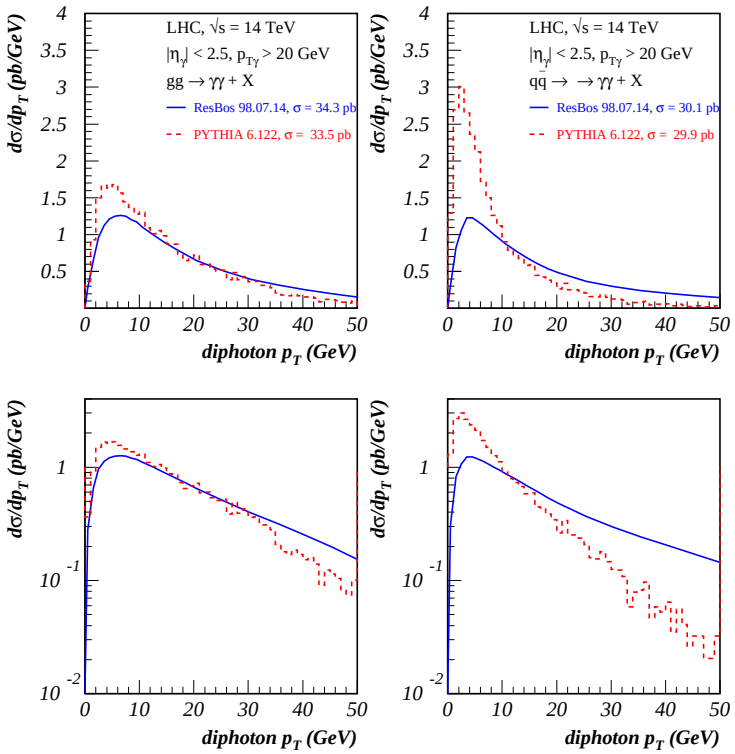

FIG. 8. A comparison of the PYTHIA and ResBos predictions for diphoton production at the LHC for the two different subprocesses, gg (left) and $q\overline{q}$ (right). Similar cuts are applied to PYTHIA and ResBos as in the ATLAS and CMS diphoton analyses. The bottom figures show the same in logarithmic scale.

calculation with parton showering cannot completely reproduce the exact matrix element calculation, without the use of matrix element corrections. The high p_T region is better reproduced if the maximum virtuality Q^2_{max} is set equal to the squared partonic center of mass energy, s, rather than m_H^2. This is equivalent to applying the parton shower to all of phase space. However, this has the consequence of depleting the low p_T region, as 'too much' showering causes events to migrate out of the peak. The appropriate scale to use in PYTHIA (or any Monte Carlo) depends on the p_T range to be probed. If matrix element information is used to constrain the behavior, the correct high p_T cross section can be obtained while still using the lower scale for showering. Thus, the incorporation of matrix element corrections to Higgs production (involving the processes $gq \to qH$,$q\overline{q} \to gH$, $gg \to gH$) is the next logical project for the Monte Carlo experts, in order to accurately describe the high p_T region.

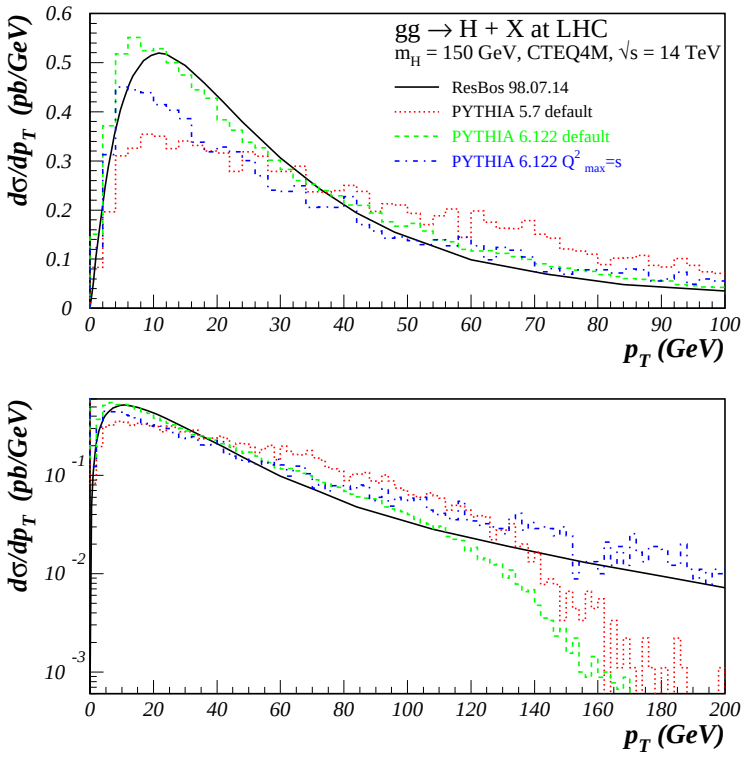

FIG. 9. A comparison of predictions for the Higgs p_T distribution at the LHC from ResBos and from two recent versions of PYTHIA. The ResBos and PYTHIA predictions have been normalized to the same area. The bottom figure shows the same in logarithmic scale.

A comparison of the two versions of PYTHIA and of ResBos is also shown in Figure 10 for the case of Higgs production (at a Higgs mass of 100 GeV) at the Tevatron with center-of-mass energy of 2.0 TeV. The same qualitative features are observed as at the LHC: the newer version of PYTHIA agrees better with ResBos in describing the low p_T shape, and there is a falloff at high p_T unless the larger virtuality is used for the parton showers. The default (rms) value of the non-perturbative k_T (0.44 GeV) was used for the PYTHIA predictions for Higgs production.

The older version of PYTHIA produces too many Higgs events at moderate p_T (in comparison to ResBos) at both the Tevatron and the LHC. Two changes have been implemented in the newer version. The first change is that a cut is placed on the combination of z and Q^2 values in a branching: $\hat{u} = Q^2 - \hat{s}(1-z) < 0$, where $\hat{s}$ refers to the subsystem of the

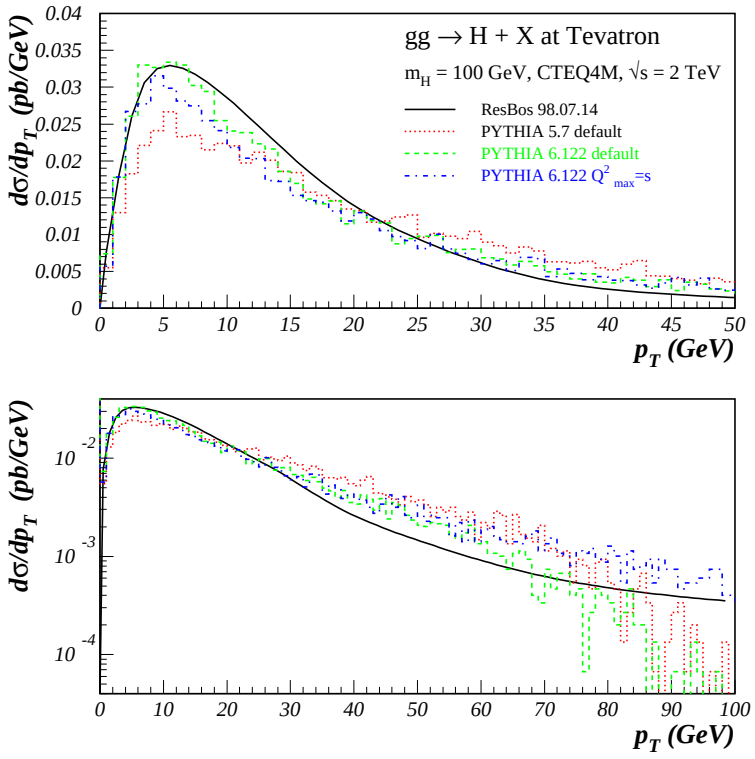

FIG. 10. A comparison of predictions for the Higgs p_T distribution at the Tevatron from ResBos and from two recent versions of PYTHIA. The ResBos and PYTHIA predictions have been normalized to the same area. The bottom figure shows the same in logarithmic scale.

hard scattering plus the shower partons considered to that point. The association with $\hat{u}$ is relevant if the branching is interpreted in terms of a $2 \to 2$ hard scattering. The corner of emissions that do not respect this requirement occurs when the Q^2 value of the space-like emitting parton is little changed and the z value of the branching is close to unity. This effect is mainly for the hardest emission (largest Q^2). The net result of this requirement is a substantial reduction in the total amount of gluon radiation [28]. [8] In the second change,

[8]Such branchings are kinematically allowed, but since matrix element corrections would assume initial state partons to have $Q^2 = 0$, a non-physical $\hat{u}$ results (and thus no possibility to impose matrix element corrections). The correct behavior is beyond the predictive power of leading log Monte Carlos.

the parameter for the minimum gluon energy emitted in space-like showers is modified by an extra factor roughly corresponding to the $1/\gamma$ factor for the boost to the hard subprocess frame [28]. The effect of this change is to increase the amount of gluon radiation. Thus, the two effects are in opposite directions but with the first effect being dominant.

This difference in the p_T distribution between the two versions of PYTHIA could have an impact on the analysis strategies for Higgs searches at the LHC. For example, for the CMS detector, the higher p_T activity associated with Higgs production in version 5.7 would have allowed for a more precise determination of the event vertex from which the Higgs (decaying into two photons) originated. Vertex pointing with the photons is not possible in the CMS barrel, and the large number of interactions occurring with high intensity running will mean a substantial probability that at least one of the interactions will have more activity than the Higgs vertex, thus leading to the assignment of the Higgs decay to the wrong vertex, and therefore a significant degradation of the $\gamma\gamma$ effective mass resolution [29]. In principle, this problem could affect the p_T distribution for all PYTHIA processes. In practice, it affects only gg initial states, due to the enhanced probability for branching with such an initial state.

As an exercise, an 80 GeV W boson and an 80 GeV Higgs were generated at the Tevatron using PYTHIA 5.7 [30]. A comparison of the distribution of values of $\hat{u}$ and the virtuality Q for the two processes indicates a greater tendency for the Higgs virtuality to be near the maximum value and for there to be a larger number of Higgs events with positive $\hat{u}$ (than W events).

VI. COMPARISON WITH HERWIG

The variation between versions 5.7 and 6.1 of PYTHIA gives an indication of the uncertainties due to the types of choices that can be made in Monte Carlos. The requirement that $\hat{u}$ be negative for all branchings is a choice rather than an absolute requirement. Perhaps the better agreement of version 6.1 with ResBos is an indication that the adoption of the $\hat{u}$ restrictions was correct. Of course, there may be other changes to PYTHIA which would also

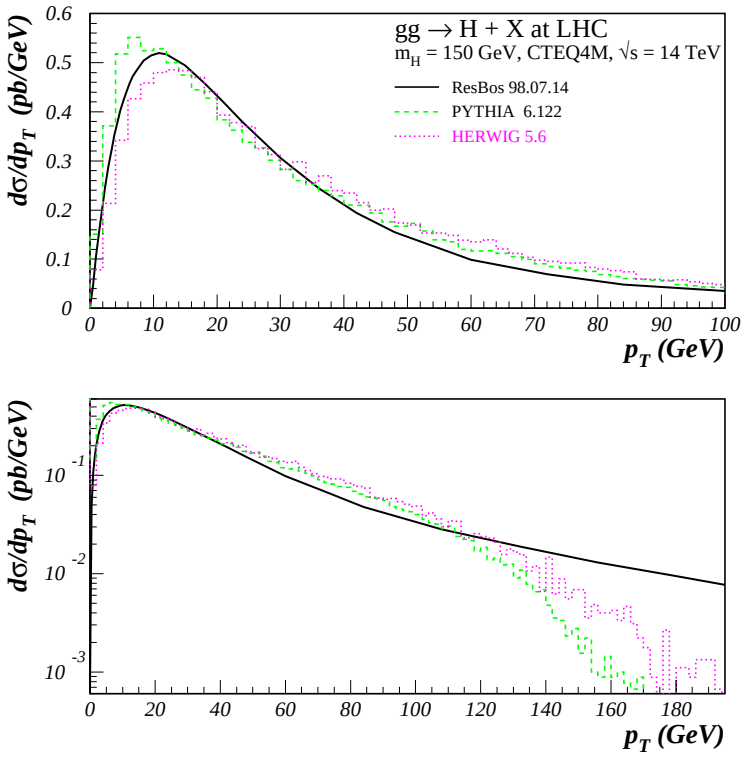

FIG. 11. A comparison of predictions for the Higgs p_T distribution at the LHC from ResBos, a recent version of PYTHIA, and HERWIG. All predictions have been normalized to the same area. The bottom figure shows the same in logarithmic scale.

lead to better agreement with ResBos for this variable.

Since there are a variety of choices that can be made in Monte Carlo implementations, it is instructive to compare the predictions for the p_T distribution for Higgs production from ResBos and PYTHIA with that from HERWIG (version 5.6, also using the CTEQ4M parton distribution functions). The HERWIG prediction is shown in Figure 11, along with the PYTHIA and ResBos predictions, all normalized to the ResBos prediction. [9] (In all cases, the CTEQ4M parton distribution was used.) The predictions from HERWIG and PYTHIA 6.1 are very similar, with the HERWIG prediction matching the ResBos shape somewhat better

[9]The normalization factors (ResBos/Monte Carlo) are 1.68 for both versions of PYTHIA, and 1.84 for HERWIG.

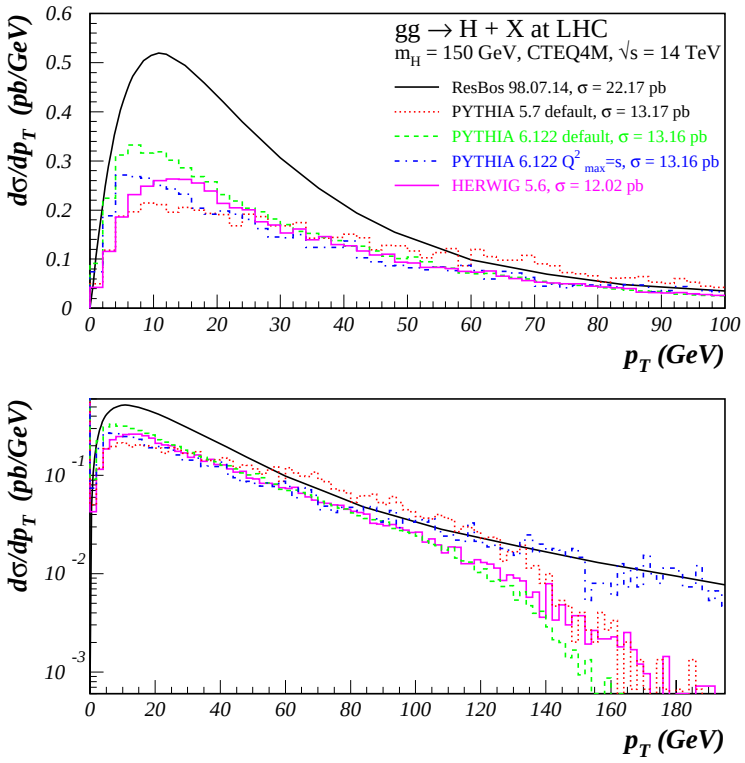

FIG. 12. A comparison of predictions for the Higgs p_T distribution at the LHC from ResBos, two recent versions of PYTHIA, and HERWIG. All predictions have their absolute normalizations. The bottom figure shows the same in logarithmic scale.

at low p_T. For reference, the absolutely normalized predictions from ResBos, PYTHIA, and HERWIG for the p_T distribution of a 150 GeV Higgs at the LHC are shown in Figure 12.

VII. NON-PERTURBATIVE K_T

A question still remains as to the appropriate input value of non-perturbative k_T in the Monte Carlos to achieve a better agreement in shape, both at the Tevatron and at the LHC. [10] In Figures 13 and 14 are shown comparisons of ResBos and PYTHIA predictions for the Higgs p_T distribution at the Tevatron and the LHC. The PYTHIA prediction (now version 6.1 alone) is shown with several values of non-perturbative k_T. Surprisingly, no difference is

[10] This has also been explored for direct photon production in Ref. [31].

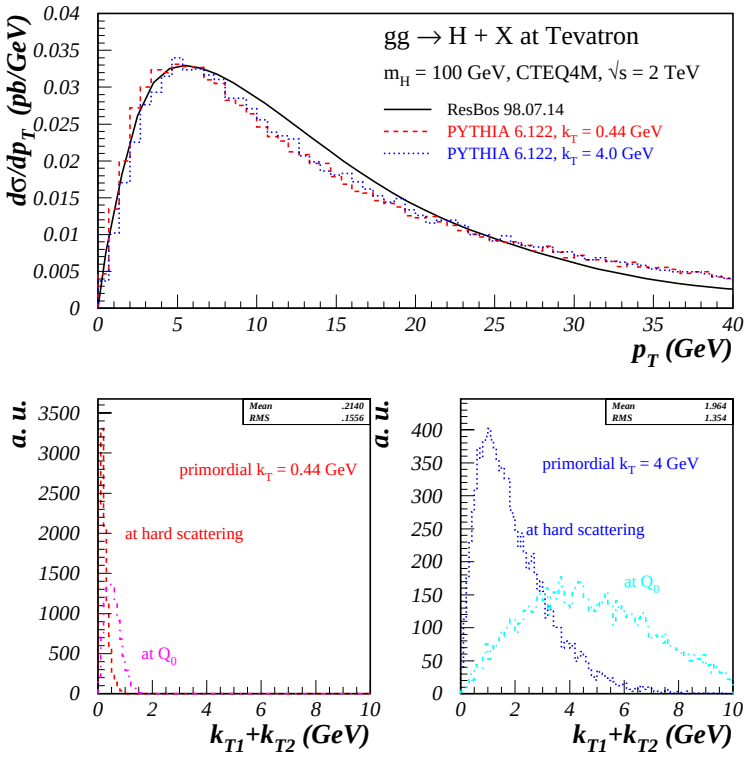

FIG. 13. (top) A comparison of the PYTHIA predictions for the p_T distribution of a 100 GeV Higgs at the Tevatron using the default (rms) non-perturbative k_T (0.44 GeV) and a larger value (4 GeV), at the initial scale Q_0 and at the hard scatter scale. Also shown is the ResBos prediction. (bottom) The vector sum of the intrinsic k_T ($\boldsymbol{k}_{T1}+\boldsymbol{k}_{T2}$) for the two initial state partons at the initial scale Q_0 and at the hard scattering scale, for the two values of primordial k_T.

observed between the predictions with the different values of k_T, with the peak in PYTHIA always being somewhat below that of ResBos. This insensitivity can be understood from the plots at the bottom of the two figures, which show the sum of the non-perturbative initial state k_T ($\boldsymbol{k}_{T1}+\boldsymbol{k}_{T2}$) at Q_0 and at the hard scatter scale Q. Most of the k_T is radiated away, with this effect being larger (as expected) at the LHC. The large gluon radiation probability from a gg initial state and the greater phase space available at the LHC lead to a stronger degradation of the non-perturbative k_T than was observed with Z^0 production at the Tevatron.

For completeness, a comparison of PYTHIA and ResBos is shown in Figure 15 for Z^0 boson

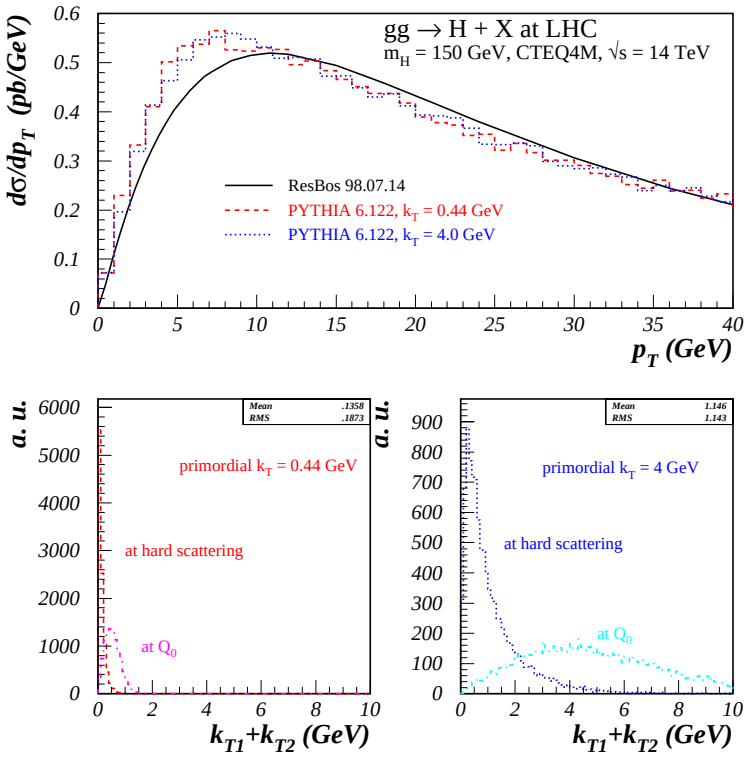

FIG. 14. (top) A comparison of the PYTHIA predictions for the p_T distribution of a 150 GeV Higgs at the LHC using the default (rms) non-perturbative k_T (0.44 GeV) and a larger value (4 GeV), at the initial scale Q_0 and at the hard scatter scale. Also shown is the ResBos prediction. (bottom) The vector sum of the intrinsic k_T ($\boldsymbol{k}_{T1}$+$\boldsymbol{k}_{T2}$) for the two initial state partons at the initial scale Q_0 and at the hard scattering scale, for the two values of primordial k_T.

production at the LHC. There are two points that are somewhat surprising. First, there is still a very strong sensitivity to the value of the non-perturbative k_T used in the smearing. Second, the best agreement with ResBos is obtained with the default value (0.44 GeV), in contrast to the 2.15 GeV needed at the Tevatron (cf. Fig1). Note again the agreement of PYTHIA with ResBos at the highest values of Z^0 p_T due to the explicit matrix element corrections applied.

The sum of the incoming parton k_T distributions, both at the scale Q_0 and at the hard scattering scale, are shown in Figure 16 for several different starting (rms) values of primordial k_T (per parton). There is substantially less radiation for a $q\overline{q}$ initial state than

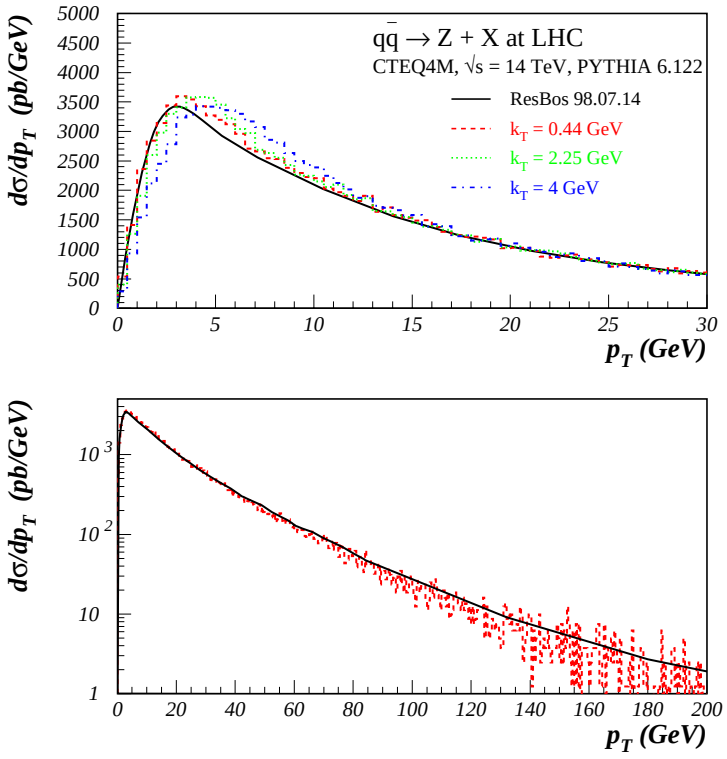

FIG. 15. A comparison of the predictions for the p_T distribution for Z^0 production at the LHC from PYTHIA and ResBos, where several values of k_T have been used to make the PYTHIA predictions.

for a gg initial state (as in the case of the Higgs), leading to a noticeable dependence of the Z^0 p_T distribution on the primordial k_T distribution.

VIII. CONCLUSIONS

An understanding of the signature for Higgs boson production at either the Tevatron or the LHC depends upon the understanding of the details of soft-gluon radiation from the initial state partons. This soft-gluon emission can be modeled either in a Monte Carlo or in an analytic resummation program, with various choices possible in both implementations. A comparison of the two approaches helps in understanding their strengths and weaknesses, and their reliability. The data from the Tevatron that either exist now, or will exist in Run 2, will be extremely useful to test both methods.

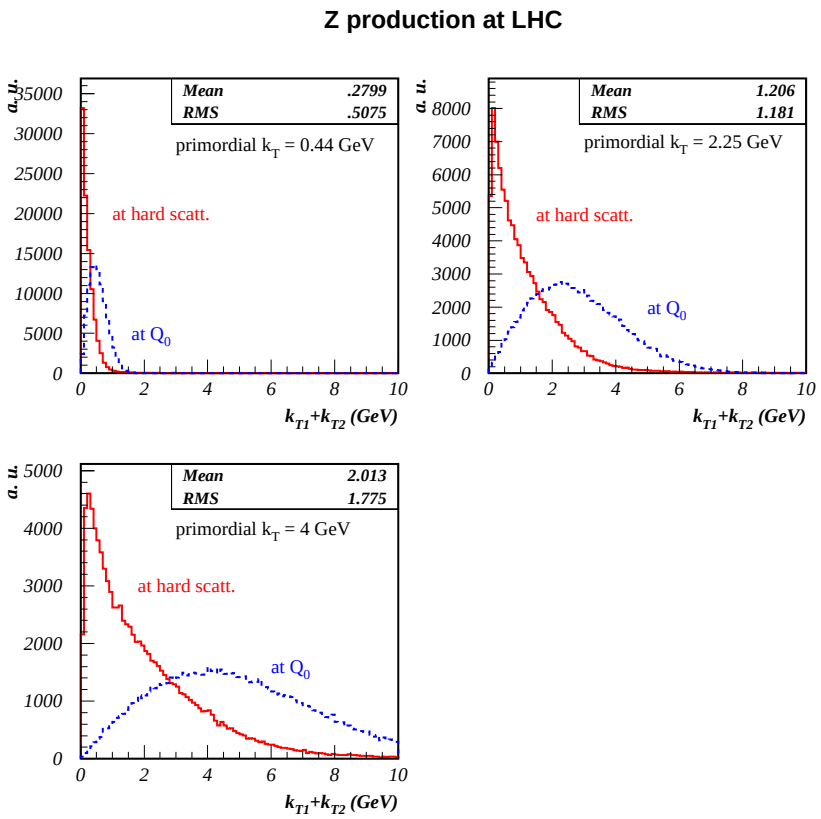

FIG. 16. A comparison of the total initial state k_T ($\boldsymbol{k}_{T1}+\boldsymbol{k}_{T2}$) distributions for Z^0 production at the LHC from PYTHIA, both at the initial scale Q_0 and at the hard scattering scale, for several (rms) values of the initial state k_T. The mean and rms numbers refer to the values at the hard scattering scale.

IX. ACKNOWLEDGEMENTS

We would like to thank Claude Charlot, Gennaro Corcella, Steve Mrenna, and Torbjörn Sjöstrand for useful conversations. We thank Willis Sakumoto for providing the figures for CDF Z^0 production, and Valeria Tano for HERWIG curves. C.B. thanks C.-P. Yuan for numerous discussions. This work was supported in part by DOE under grant DE-FG-03-94ER40833, and by NSF under grant PHY-9901946.

[1] Proceedings of *Physics at Run 2: Workshop on Supersymmetry and Higgs*, Summary Meeting, Batavia, IL, Nov. 19-21, 1998.

[2] S. Abdullin, M. Dubinin, V. Ilyin, D. Kovalenko, V. Savrin and N. Stepanov, Phys. Lett. **B431**, 410 (1998) [hep-ph/9805341].

[3] C. Balázs, P. Nadolsky, C. Schmidt and C.–P. Yuan, hep-ph/9905551.

[4] C. Balázs and C. P. Yuan, Phys. Rev. **D59**, 114007 (1999) [hep-ph/9810319].

[5] C. Balázs and C.–P. Yuan, hep-ph/0001103.

[6] T. Sjöstrand, Phys. Lett. **B157**, 321 (1985).

[7] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994).

[8] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, Comput. Phys. Commun. **67**, 465 (1992).

[9] F.E. Paige, S.D. Protopescu, H. Baer and X. Tata, hep-ph/9810440.

[10] ATLAS Detector and Physics Performance Technical Design Report, CERN/LHCC/99-14.

[11] J.C. Collins and D.E. Soper, Phys. Rev. Lett. **48**, 655 (1982); Nucl. Phys. **B193**, 381 (1981), **B213**, 545(E) (1983); **B197**, 446 (1982);

J.C. Collins, D.E. Soper and G. Sterman, Nucl. Phys. **B250**, 199 (1985).

For the application of the CSS formalism to Z^0 boson and diphoton production see: [4,13] and C. Balázs, E.L. Berger, S. Mrenna and C.–P. Yuan, Phys. Rev. **D57**, 6934 (1998) [hep-ph/9712471];

C. Balázs, Ph.D. thesis, Michigan State University (1999), hep-ph/9906422.

[12] C. Balázs, J.C. Collins and D.E. Soper, "Generalized factorization and resummation", in the proceedings of the *Workshop on Physics at TeV Colliders*, Les Houches, France, June 8-18, 1999.

[13] C. Balázs and C.–P. Yuan, Phys. Rev. **D56**, 5558 (1997) [hep-ph/9704258].

[14] M. Spira, hep-ph/9711394.

[15] C. Schmidt, private communication.

[16] S. Catani and B.R. Webber, Nucl. Phys. **B349**, 635 (1991).

[17] G. Miu and T. Sjöstrand, Phys. Lett. **B449**, 313 (1999).

[18] G. Corcella and M.H. Seymour, RAL-TR-1999-051 and hep-ph/9908388.

[19] H. Baer and M. H. Reno, Phys. Rev. **D44**, 3375 (1991); **D45**, 1503 (1992).

[20] S. Mrenna, hep-ph/9902471.

[21] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **84**, 845 (2000) [hep-ex/0001021].

[22] G. Corcella, talk at the LHC workshop, October 1999.

[23] H.L. Lai, J. Huston, S. Kuhlmann, F. Olness, J. Owens, D. Soper, W.K. Tung, and H. Weerts, Phys. Rev. **D55**, 1280 (1997).

[24] G. Ladinsky, C.–P. Yuan, Phys. Rev. **D50**, 4239 (1994).

[25] F. Abe *et al.*, Phys. Rev. Lett. **70**, 2232 (1993);

T. Takano, Ph.D. thesis, U. Tsukuba (1998);

CDF Collaboration, paper in preparation.

[26] W. Chen, Ph.D. thesis, State University of New York, Stony Brook, NY (1997);

D0 Collaboration, paper in preparation.

[27] P. Aurenche, A. Douri, R. Baier, and M. Fontannaz, Z. Phys. **C29**, 423 (1985);

B. Bailey, J. Owens, and J. Ohnemus, Phys. Rev. **D46**, 2018 (1992);

T. Binoth, J.P. Guillet, E. Pilon, and M. Werlen, hep-ph/9911340.

[28] `PYTHIA` manual update for version 6.1, `http://www.thep.lu.se/~torbjorn/Pythia.html`.

[29] D. Denegri, private communication.

[30] S. Mrenna, talk at the *Physics at Run 2: Workshop on Supersymmetry and Higgs*, Summary Meeting, Batavia, IL, Nov. 19-21, 1998;

C. Balázs, J. Huston, S. Mrenna, and I. Puljak, in the proceedings of *Physics at Run 2: Workshop on Supersymmetry and Higgs*, Summary Meeting, Batavia, IL, Nov. 19-21, 1998.

[31] H. Baer and M. H. Reno, Phys. Rev. **D54**, 2017 (1996) [hep-ph/9603209].